FACEBOOK POST ANALYSIS FOR A CELEBRITY PAGE

Lovin Babu Thomas

Department of Computer Science, CHRIST (Deemed to be University),
Bengaluru, India

Joy Paulose

Department of Computer Science, CHRIST (Deemed to be University),
Bengaluru, India

ABSTRACT

Social Media is a subject that is large and extremely talked about in the present day. The expanding accessibility of the web and the development of social media brought in a necessity for the analysis of data based on a Facebook page. The number of information generated by social media platforms are expanding exponentially. The best approach to understand data of this magnitude lies in mining and analysis of such big data. The present focus is on the analysis of a Facebook Fan Page. This research work concentrates on the examination of the number of likes, post shares and correlation of posts and comments on various topics, fields or areas of a Facebook page. 'Like' is not the only way of rating a post nowadays. There are also recently added features like emoticons which display emotions such as 'love', 'wow', 'angry', 'sad', etc., to rate a post. The objective of this paper is to extract the data available and to perform analysis on a cricket fan page. Python is used to extract and analyse the information.

Key words: Social media, Facebook, Emoticons, Data extraction, Python.

1. INTRODUCTION

In today's world, social networking is a common word that is discussed quite a lot in many technical circles. Web-based social networking is the new pattern, a trend that is focused on promoting Information Technology on various levels. This phenomenon is not something that recently came into vogue, yet with the steady advancement in internet technology, everyone's acquaintance with these tools are expanding. The analysis of social networking sites stands out amongst the most prevalent research topics in today's world. With the quick development of informal communities, some new organizations make a page on Facebook and post news or topics that they need to share with their peers.

Facebook is one of the top social networking site in the world. Social networking is an interconnected stage that establishes a social connection between individuals with common interest and paves the way to associate with other interdisciplinary regions like business, sports, politics, and financial institutions. Facebook has nearly 2 billion active users on a monthly basis and more than 20 million people become followers of one or the other Facebook pages every day. In addition to this, the fact that there are over 3 million pages with varying fields of interest is intriguing. In the field of marketing too, many changes have happened after the development of the social network. Established companies make fan pages for their customers to promote their businesses on their respective business pages by posting and sponsoring different posts so that it reaches as many people as possible. The post reaches can be analysed by the number of likes, reactions, comments and shares.

The positive aspect of social networking as an 'extra showcasing channel' is that it can be utilized to convey messages on a global level and at the same time, it also enhances correspondence towards people at an individual level. A 'Facebook Wall' is the focal point of every user profile for sharing publicly, new media like photographs, videos, notes and most importantly, thoughts or feelings that a user might be indulging in at that particular post.

The purpose of this paper is to analyse celebrity fan pages on Facebook and to see how extensively the uploaded posts can reach out, to comprehend the total number of likes and reactions that these posts get and to also understand what could be done to increase the reach of a particular post. The language used for this research work is Python. This research study focuses on a cricket fan page on Facebook and goes through various posts and sees how far each post has reached. The total count of likes, comments and reactions can also be fetched. A post cannot be classified as good or bad by just focusing on the number of likes it generates. There are emoticons which can, to some extent, determine whether a post is good or bad. The types of emoticons available are 'love', 'wow', 'haha', 'sad' and 'angry'. 'Love', 'wow' and 'haha' can be categorized as positive and 'sad', 'angry' can be classified as negative. If a celebrity is seen to be sharing his own photos, then the reach of those photos are more, when compared to other posts. The reason behind the need for this analysis is to see what topics people are more interested in, how far a post can reach and how much of an impact it can have on the audience. Predictions can be made based on the analysis and comparisons.

2. BACKGROUND

Irena Pletikosa Cvijikj [1] focused on the field of marketing. There were many changes that the global market faced after the spontaneous growth of the social network. The main advantage of promoting products through social media is that, interactions can be global and that it reaches to a large number of individuals. The objective of this work is to assess the impact of the three fundamental qualities of arbitrator posts on the level of user interactions in a sponsored Facebook brand page.

The interaction made by a user is measured by seeing the number of comments on each post, the number of likes on each post and the duration of interactions on a large dataset. They have implemented Kruskal–Wallis non-parametric test for analysis. All three types of posts have different levels of interaction and the posts containing status tended to have more comments. The videos received a lot more likes than links and photos, and these media ended up having lesser amounts of interaction. Interactions are more on photos and it was followed by statuses, videos, links and posts. Based on the observations, different moderator posts having different levels of activities have been identified.

Antonio Teixeira, et al. [2] in their research work, concludes that among social networking sites, Facebook is the best platform for understanding emotions. Recently, Facebook introduced reactions, which can be used to express feelings toward a particular post. Reactions are love, haha, like, wow, sad and anger. People can express sentiments through comments or emoticons. Emoticons are divided into different categories like positive, extremely positive, negative, extremely negative and neutral. These days, user interaction with brand pages on Facebook is getting more relevance. This research work shows how to extract, pre-process and prepare data collected from Facebook for sentiment analysis.

Seyed Morteza Ghavami, et al. [3] propose that the users have different behaviours. Some are talkative while others are not. Personality plays a vital role in a recommender system. This paper deals with relations between Facebook users with respect to behaviour and personality. Two graphs have been plotted namely PostLikeGraph and CommentLikeGraph, data collection is from the whose post. PostLikeGraph is based on the analysis of who likes public post. CommentLikeGraph is based on who likes whose comment. From this outcome, preparing a classifier by means of user likes could also predict user personality characteristics.

Essi Pöyry, et al. [4] discusses that Facebook pages are used as a marketing tool by branded companies. The interactions with these pages and customer intentions to purchase these products are not clear. The usage behaviour of a community and the intentions behind the purchase of products they are discussing here. The findings result in understanding user interactions and behaviours in branded. Facebook pages, is also a reason to buy the host company's product.

Chang-Ling Hsu, et al. [5] focuses on Big Data-Fuzzy Grey Relational Analysis (BD-FGRA) algorithm to select top-N posts that satisfies all customers' specified selection method, the criteria of ranking and so on. It ranks posts of candidates with this novel algorithm. Compared to the other methods BD-FGRA algorithm has better performance and a very convincing execution time and precision.

Kuan-Cheng Lin, et al. [6] discusses the analysis of user sentiments from corporation-run fan pages on Facebook. The data set for this study is collected through likes, shares and comments from the page. The aim is to find the topics and sentiments in a given Facebook page. The analysis of sentiments and information extraction together results in finding the most interesting events on the Facebook fan page. Hsin-Ying Wu, et al. [7] has presented a research work on text mining technique over Facebook pages. Some of the pages receive good feedback, while others receive bad feedback. Text mining performed on these pages leads to better management of the pages. Data collection is through Facebook by collecting consumer feedback. Extraction of key phrases is collected from the Chinese language and Chinese Knowledge and Information Processing (CKIP) is applied for extraction of data. Clustering is used to generate the critical points of customer needs and to find solutions for their needs.

Imen Moalla, et al. [8] present the information about data warehouses, the design and the multidimensional schema concepts. The amount of data produced on social media is extremely large. Therefore, the analysis and storage of this huge chunk of data in a data warehouse is very important. Thus the idea of data warehouse construction from social media came into the picture. This model can estimate the success of the products and also to improve the success rates for a product that is in existence.

Rehab M. Duwairi, et al. [9] has presented a research work which explains, RUM: a data extraction tool which is used to extract data and content from Facebook. The extracted data are further saved and analysed.

Sudipto Shankar Dasgupta, et al. [10] present the information about the techniques for collecting the feedbacks regarding brand companies is by social network. Therefore, it is equally important to mine the unstructured information from social networks. The use of open source technologies provides easy accessibility, high quality software and abundant support to mine unstructured data. This paper focuses on the usage of open source technologies to mine data.

3. METHODOLOGY

The objective of this work is to extract the data from a Facebook page and analyse the data to understand the nature of the expression of the respondents. To achieve this, the following methodology is proposed which comprises of the following four steps:

1. Data collection: In this step, data collection is from the Social Media   Platforms. Here, Facebook is the platform used to access data. The data are in an unstructured form.

2. Data preparation: Removing all non-textual contents of the data collected.

3. using the data analyze and plot a graph.

4. Finally, compare different posts and see how far it reaches people with the help of graphs.

The main step in this process is the extraction of data from Facebook. The first step is to create a Facebook dashboard and to configure our Facebook account on the dashboard. Facebook dashboard is nothing but an application setting dashboard which is used for better results. An app id must be added under MyApps in dashboard settings. To fetch data, access tokens should be generated, which happens to be on the Facebook dashboard. Application Program Interface(API), a tool for building application software. Access token is a key which provides temporary and secure access to Facebook APIs. 'Temporary' is mentioned because access tokens will expire after a period of time, so it should be generated again to fetch data from Facebook. Graph API is a low-level hypertext transfer protocol API which allows to perform various operations and one of them is the extraction of data from Facebook. The simplest method to fetch data is by using APIs.

The tool used is Python, for extraction and collection of data sets. The first step is to import the library functions. It includes importing Facebook data. Import urllib3, as it contains some standard library functions which are not available in Python standard library. Other import functions include requests, re, JSON.
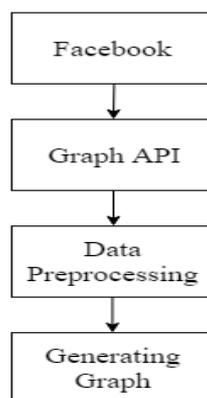
Figure 1 Steps comprising of the whole process.

4. EXPERIMENTAL RESULT

In this research, a cricket celebrity's Facebook page is taken for analysis to see the post's reach, the reasons for a post to get more likes and how a post can reach more people. Code for extracting the data from Facebook is given below:

```
graph=facebook.GraphAPI(access_token=" ")
posts=graph.get_connections(id=" ",connection_name="likes", summary='true')
posts['summary']['total_count']
```

In the above code, people who liked a particular post and the total likes that post received are identified. Creating Facebook graph API objects and then passing access tokens as its parameters. Access token is a key to access data or information from the Facebook platform. Here, 'id' is the identification number of a particular post. Using the id, we try connecting these graphs to the Facebook server and getting names of people who liked a particular post. Finally, collecting the total like's summary. This is the simplest way of collecting data using python.

In Fig 2, five different posts are analysed and the total likes of each post are counted and analysis is based on the collected data
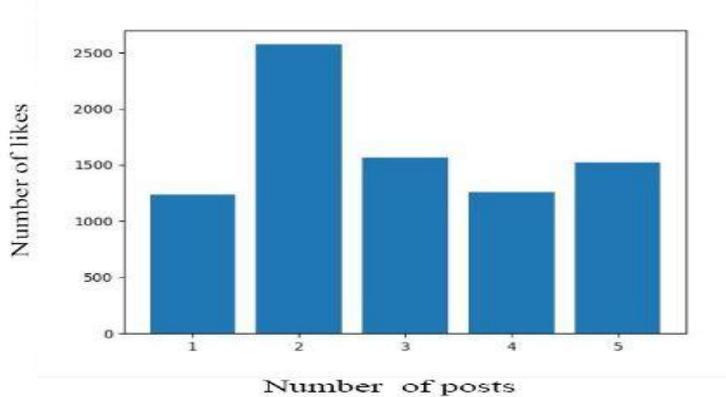

Figure 2 Likes of five posts

Here, focus is on analysing a cricket celebrity fan page. From this page, various posts are analysed to understand what the reasons are for different post reachability and various other factors. Fig. 3 shows the analysis of 10 posts of a celebrity and each line data represents reactions to the post
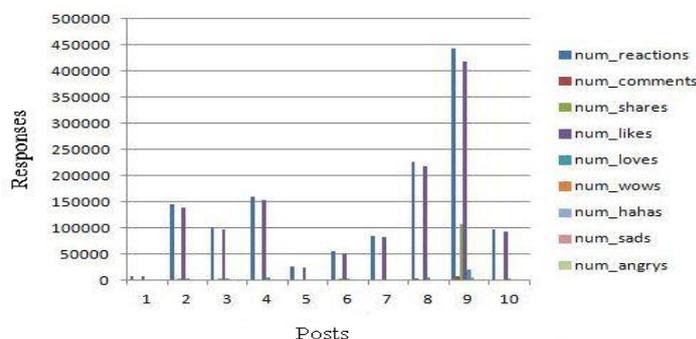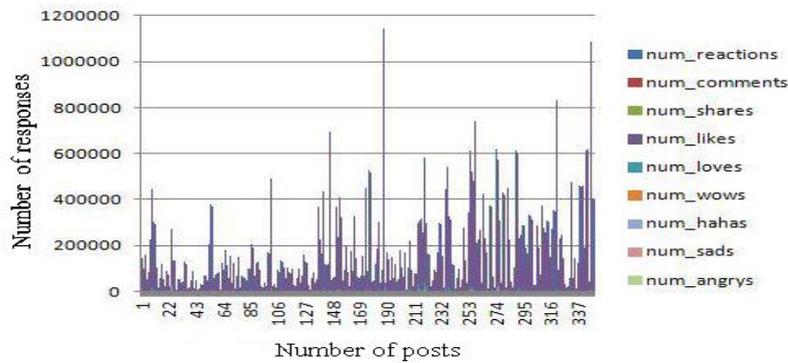

Figure 3 Analysis of a sample of 10 posts.

It also includes the total number of comments, shares and reactions. The post can be classified as positive if the reactions like haha, wow and love are more than the other emoticons and if it isn't the case, then it can be classified to be negative. Fig. 4 shows the graph, which presents 350 random posts from Sachin's Facebook page.

Figure 4 Analysis of 350 different posts.

The posts include photos, videos, links and statuses. Some of the posts are shared by many while the others are shared less in comparison. 'Like' is not the only way to analyse whether a post is positive or negative. The reactions are an easier way to analyse a post. If positive reactions (love, haha, wow) are more, the post is a positive post. In Fig. 5, a celebrity's random posts data is collected and analyzed.
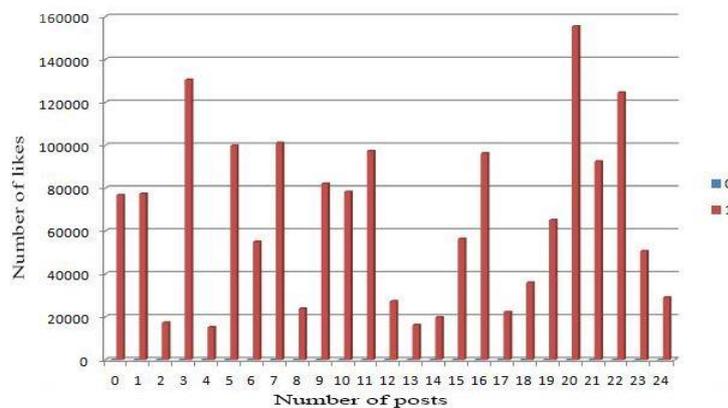
Figure 5 Likes of 25 different photos

The limit is given in such a way that 25 photos are analyzed and the data is collected. Some of the posts get more than one lakh likes and some post likes are less compared to other posts. The result after analysis is that, the post with photos of celebrity is reaching more audience and getting more likes. Posts about upcoming events are also getting more reach. If any cricket match is going to happen in the coming days, the interactions and the reach of the post is more comparable to other posts. In this way, analysis of match can be made to reach more people and audience's interest towards a particular match or event based on the post analysis.

4. CONCLUSIONS

In this paper, analysis is made based on the likes, comments and reactions of the Facebook posts. The motive of this research is to extract and prepare data from a Facebook cricket fan page. The objective is to see the reasons behind a post reaching more people and then predictions made by the analysis. A post can be categorized into positive or negative based on the reactions. If a post related to any cricketing event is happening in the future, then by analyzing the likes and interactions, the audience interest in that event and predictions can be made by making use of the analysis.

In this paper, a small data set is used but in the future, larger data set can be used to predict with more accuracy. So, having many number of posts will give better results.

REFERENCES

[1]   Irena Pletikosa Cvijikj, "The Effect of Post Type, Category and Posting Day on User Interaction Level on Facebook", IEEE 2011.

[2]   Antonio Teixeira,"Data extraction and preparation to perform a sentiment analysis using open source tools " , IEEE 2015.

[3]   Seyed Morteza Ghavami,"Facebook User's Like Behavior Can Reveal Personality", IEEE 2015.

[4]   Essi Pöyry. "The Power of 'Like' – Interpreting Usage Behaviors in Company-Hosted Facebook Pages",IEEE 2012.

[5]   Chang-Ling Hsu, "A Big Data Fuzzy Grey Relational Analysis Method for Social Post Retrieval– with an Application to Facebook's News Groups ", IEEE 2017.

[6]   Kuan-Cheng Lin, "Mining the User Clusters on Facebook Fan Pages Based on Topic and Sentiment Analysis ", IEEE 2014.

[7]   Hsin-Ying Wu, "Understanding Customers Using Facebook Pages: Data Mining Users Feedback Using Text Analysis ", IEEE 2014.

[8]   Imen Moalla, "Data warehouse design from social media for opinion analysis: the case of Facebook and Twitter", IEEE 2016.

[9]   Rehab M. Duwairi, "RUM Extractor: A Facebook Extractor for Data Analysis ", IEEE 2015.

[10]   Sudipto Shankar Dasgupta,"Sentiment Analysis of Facebook Data using Hadoop based Open Source Technologies", IEEE 2015.

[11]   https://www.facebook.com/SachinTendulkar/Sachin's Facebook page

[12]   Dr. Rashmi Gopinathan and Dr. Sapna Suri, A Study of Awareness and Usage of Social Media in HR Practices and It's Impact on Organizations Especially with Reference To Pharma Industry. International Journal of Management, 7(5), 2016, pp. 164–172.

[13]   Roby Setiadi, Ellin Herlinas, Bara Bagus Priharto Wike Pertiwi and Mohamad Djadjuli, Adaptive Management Strategies on Advances of Information Technology: A Study of the Importance of Social Media Marketing and Customer Database, International Journal of Civil Engineering and Technology, 8(11), 2017, pp. 1050–1058

[14]   Unnati Patel and Dr. Govind B Dave, An Empirical Study of Advertisements of Products on Social Media Platforms, Analysing The Sectoral Predictability of Risk and Return In India . International Journal of Management, 7 (4), 201 6 , pp. 195 – 207 .

[15]   Esther Hepziba. R and Dr. Florence John, an Exploratory Study on the Influence of Social Media Marketing Strategies on Customer Engagement. International Journal of Marketing and Human Resource Management, 8(3), 2017, pp. 1–8.

[16] Prabha Kiran and Dr. S Vasantha, Antecedents of Social Media Information Search towards Purchase Decision. International Journal of Civil Engineering and Technology, 8(6), 2017, pp. 647–653.

[17] Dr. Meenakshi Tomar, Dr. Krishan K Pandey and Dr. Devendra K Punia, Customer Engagement on Social Media: Understanding the Functional Perspectives of Social Media Towards Customer Engagement. International Journal of Marketing and Human Resource Management , 7(3), 2016, pp. 60–63.

[18] Rajapriyan and N. Kumar, Load -Balanced Migration of Social Media to Content Clouds . International Journal of Civil Engineering and Technology , 8(3), 2017 , pp. 720 –728